Online Appendix: A Simple Model of Searching for an Effective Teacher

In this appendix, we outline a simple model of using imperfect estimates of teacher effectiveness to screen out ineffective teachers and maximize student achievement. Our model is analogous to standard models of job search in which there is learning about productivity on the job (Jovanovic, 1979) and builds on a model developed in more detail in Kane and Staiger (2005). Instead of a worker searching for the most productive job, we have a principal searching for the most productive teacher. In the language of search models, we assume teachers are an experience good (Mortensen, 1986) – principals can learn only so much at the time of hire, and must learn more about teacher productivity by observing performance on the job. Thus, the principal draws teachers from the applicant pool, observes noisy signals over time about teacher productivity, and decides whether to dismiss unproductive teachers and start the process over again.

We assume that the process by which a principal searches for an effective teacher is fairly simple. First, the principal collects applications and gathers information about each applicant. Based on this information (the "pre-hire" signal), she chooses the most promising candidates to fill the available vacancies. Each year she gathers additional information on each new hire's performance in the classroom (the "on-the-job" signal). The principal may dismiss a teacher at the end of each year until the teacher reaches tenure (usually after the 3rd year). If she chooses to dismiss a teacher, or if a teacher chooses to leave for other exogenous reasons, she must start the process over again. Teacher turnover is costly because of the time and effort involved in dismissing and recruiting a new teacher, and because replacement teachers will have no prior experience. Ultimately, the principal tries to manage this process in a way that maximizes average teacher effectiveness in her school.

The General Model

Suppose that teacher effectiveness (μ) is normally distributed in the population so that:

$$(1) \qquad \mu \sim N\left(0, \sigma_\mu^2\right)$$

The principal observes a noisy pre-hire signal of teacher effectiveness (Y_0) that is normally distributed with mean μ:

$$(2) \qquad Y_0 \sim N\left(\mu, \sigma_0^2\right)$$

Thus, the reliability of the pre-hire signal is $\sigma_\mu^2 / \left(\sigma_\mu^2 + \sigma_0^2\right)$, the ratio of signal variance to the sum of signal variance and noise variance.

We assume that the number of teachers applying each year is equal to 10 times the natural attrition rate (if there were no dismissal), which is in line with evidence from LA and NYC. In other words, in the state of the world where no teachers are dismissed—very close to its current state—there are 10 applicants for each vacancy. The number of teachers hired from this pool each year must be equal to the number of teachers leaving through natural attrition plus the number dismissed. We assume that teachers work for at most 30 years, and, if not dismissed, have an exogenous annual turnover rate δ.

Teacher effectiveness on the job is not observed directly, and the principal must search for effective teachers based on a normally distributed annual performance signal (Y_t), which depends on both the teacher's effectiveness (μ) and a return to experience (β_t):

$$(3) \qquad Y_t \sim N\left(\mu + \beta_t, \sigma_\varepsilon^2\right)$$

We normalize the effectiveness of an experienced teacher to be zero, and let $\beta_t < 0$ (t = 1, 2) represent the differential between an inexperienced and experienced teacher. Additional hiring costs can be included by making β_1 more negative.

The principal can dismiss a teacher at the end of each year up until the time of tenure (T). The principal's objective is to maximize the average effectiveness over all teachers in steady state, which is equivalent to maximizing expected productivity in a search model with no discounting.

The optimal strategy for this type of search model has the reservation property: at the end of each year (t), the principal dismisses a teacher if their expected effectiveness given the information to date ($E(\mu|Y_0,.., Y_t)$) lies below a reservation value (r_t), where $E(\mu|Y_0,.., Y_t)$ is derived from the normal learning model.

The reservation value rises with time on the job, because the option value of waiting to dismiss a teacher declines as the principal accumulates better information over time. In other words, to avoid unnecessary turnover the principal may choose to wait a year before dismissing a teacher who she believes is "below the bar," so long as there is a reasonable chance that her beliefs could change. Thus, the principal dismisses teachers whose expected effectiveness lies below a bar that increases with teacher experience. Overall, the principal must set the bar to trade off the short-term cost of replacing an experienced teacher with a rookie against the long-term benefit of selecting only the most effective teachers.

There is no simple closed form solution for calculating the reservation value. However, the optimal reservation value depends on the underlying parameters of the model: the variation in performance across teachers (σ_μ^2), the strength of the pre-hire and on-the-job signals ($\sigma_0^2, \sigma_\varepsilon^2$), the return to experience (β_t), the number of years before tenure (T), the exogenous turnover rate (δ), and the size of the applicant pool relative to the exogenous turnover rate. We used evidence on these key underlying parameters from New York and Los Angeles to simulate data, and then solved for the optimal reservation values numerically.

An Illustrative Example

A simplified version of this general model yields an analytical solution that illustrates the tradeoffs facing a principal. Suppose that a teacher's classroom performance each year (Y_t) is the sum of a persistent teacher effect (μ), an error term that is independent across years (ε_t), and a negative "rookie" effect ($\beta<0$) in the first year only (t=1), where both μ and ε_t are normally distributed with mean zero. Further assume that there is no pre-hire signal, so that each new hire is a random draw from the teacher distribution. Finally, suppose that new hires must be either dismissed or tenured at the end of their first year (*T*=1). This simplifies the solution by eliminating any option value of waiting.

The principal will grant tenure if a teacher's classroom performance in the first year exceeds a reservation value ($Y_1 > r_1$). The cut-off value for tenure after 1 year (r_1) is chosen to maximize the average productivity of the entire workforce. The workforce consists of two groups of workers: rookies in their first year of teaching, whose expected performance is just β, and teachers who survived the tenure cut-off, whose expected performance is $E(\mu \mid Y_1 > r_1)$. Therefore, the average productivity of the workforce ($\bar{Y}$) is equal to:

$$\text{(4)} \qquad \bar{Y} = \pi\beta + (1-\pi)E(\mu \mid Y_1 > r_1)$$

Where the proportion of rookies in the workforce in steady state (π) is given by:

$$\text{(5)} \qquad \pi = \frac{1}{1 + \sum_{t=2}^{T} (1-\delta)^{t-1} Pr(Y_1 > r_1)}$$

Thus, π is an increasing function of both the exogenous turnover rate and the tenure cutoff below which rookies are dismissed. Thus, raising the cutoff increases the expected productivity of teachers reaching tenure, but at the cost of raising the proportion of the workforce who are rookies.

Maximizing $\bar{Y}$ with respect to r_1 yields the following simple first-order condition determining the choice of the optimal value of r_1:

$$\text{(6)} \qquad E(\mu \mid Y_1 = r_1) = \bar{Y}$$

The above expression has a fairly straightforward interpretation. The expression on the left is the productivity of the *marginal* teacher, whose performance was at the reservation value r_1. The expression on the right is the productivity of the *average* teacher (including both tenured teachers and rookies). The principal sets the cut-off, r_1, where the productivity of the *marginal* teacher is equal to the productivity of the *average* teacher. In other words, this decision rule tells principals to keep only the rookies who are expected to be better than the average teacher. This result is analogous to the usual result that average costs are minimized at the point where marginal cost equals average cost.

The above first order condition, in combination with the definition of average productivity in Equation 4, has a number of implications for the determinants of the cut-off level of performance required for tenure. First, a more negative rookie effect (β) lowers the average productivity of the workforce, which in turn lowers the optimal reservation value. Put simply, the value of experience raises the cost of dismissing experienced teachers. Similarly, a high exogenous turnover rate raises the fraction of rookies in the workforce (π) and lowers the average productivity of the workforce, which again lowers the optimal cutoff for tenure. There is less benefit to giving tenure to highly effective teachers if they do not stay long. Finally, low variance in the teacher effect (μ) lowers the benefit of selection, and high variance in the error with which productivity is measured (ε_t) makes it more difficult to select highly effective teachers, both of which lower the optimal cutoff for tenure. There is little reason to be selective if the performance data (Y_1) cannot reliably identify important productivity differences between teachers.